

IN THE CLAIMS:

The status of the claims is noted below.

1. (Currently Amended) A signal processing method for detecting and analyzing a pattern reflecting the semantics of the content of a signal, the method comprising steps of:

extracting, from a segment consisting of a sequence of consecutive frames forming together the signal, at least one feature which characterizes the properties of the segment;

calculating, using the extracted feature, a criterion for measurement of a similarity between a the pair of segments for every extracted feature and measuring a similarity between [[a]] the pair of segments according to the similarity measurement criterion; and

detecting, according to the feature and similarity measurement criterion, two of the segments, whose mutual time gap is within a predetermined temporal threshold and mutual dissimilarity is less than a predetermined dissimilarity threshold, and grouping the segments into a scene consisting of a sequence of temporally consecutive segments reflecting the semantics of the signal content.

2. (Original) The method as set forth in Claim 1, wherein the signal is at least one of visual and audio signals included in a video data.

3. (Original) The method as set forth in Claim 1, wherein at the feature extracting step, a single statistic central value of the plurality of features at different time points in a single segment is selected for extraction.

4. (Original) The method as set forth in Claim 1, wherein a statistic value of the similarity between a plurality of segment pairs is used to determine the dissimilarity threshold.
5. (Original) The method as set forth in Claim 1, wherein of the segments, more than at least one segment which could not have been grouped into a scene at the grouping step are grouped into a single scene.
6. (Original) The method as set forth in Claim 1, wherein a result of scene detection from arbitrary features acquired at the grouping step and more than at least one result of scene detection for features different from the arbitrary ones, are combined together.
7. (Original) The method as set forth in Claim 2, wherein more than at least one result of scene detection from the video signal acquired at the grouping step and more than at least one result of scene detection from the audio signal acquired at the grouping step, are combined together.
8. (Currently Amended) A video signal processor apparatus for detecting and analyzing a visual and/or audio pattern reflecting the semantics of the content of a supplied video signal, the apparatus comprising:

means for extracting, from a visual and/or audio segment consisting of a sequence of consecutive visual and/or audio frames forming together the video signal, at least one feature which characterizes the properties of the visual and/or audio segment;

means for calculating, using the extracted feature, a criterion for measurement of a similarity between a pair of visual segments and/or audio segments for every extracted feature and measuring a similarity between [[a]] the pair of visual segments and/or audio segments according to the similarity measurement criterion; and

means for detecting, according to the feature and similarity measurement criterion, two of the visual segments and/or audio segments, whose mutual time gap is within a predetermined temporal threshold and mutual dissimilarity is less than a predetermined dissimilarity threshold, and grouping the visual segments and/or audio segments into a scene consisting of a sequence of temporally consecutive visual segments and/or audio segments reflecting the semantics of the video signal content.

9. (Original) The apparatus as set forth in Claim 8, wherein the feature extracting means selects, for extraction, a single statistic central value of the plurality of features at different time points in a single visual and/or audio segment.

10. (Original) The apparatus as set forth in Claim 8, wherein a statistic value of the similarity between a plurality of visual and/or audio segment pairs is used to determine the dissimilarity threshold.

11. (Original) The apparatus as set forth in Claim 8, wherein of the visual and/or audio segments, more than at least one visual and/or audio segment which could not have been grouped into a scene by the grouping means are grouped into a single scene.

12. (Original) The apparatus as set forth in Claim 8, wherein a result of scene detection for arbitrary features acquired by the grouping means and more than at least one result of scene detection for features different from the arbitrary ones, are combined together.

13. (Original) The apparatus as set forth in Claim 8, wherein more than at least one result of scene detection from the visual signal of the video signal acquired by the grouping means and more than at least one result of scene detection from the audio signal of the video signal acquired by the grouping means, are combined together.